

CLAIMS

1 1. (currently amended) A method for synthesizing an auditory scene, comprising the steps
2 of:
3 (a) dividing an input audio signal into a plurality of different frequency bands; and
4 (b) applying two or more different sets of one or more spatial parameters to two or more of
5 the different frequency bands in the input audio signal to generate two or more synthesized audio signals
6 of the auditory scene, wherein, for at least one of the sets of one or more spatial parameters, at least one
7 of the spatial parameters corresponds to a combination of two or more different audio sources in the
8 auditory scene that takes into account relative dominance of the two or more different audio sources in
9 the auditory scene.

1 2. (previously pending) The invention of claim 1, wherein:
2 the input audio signal corresponds to a combination of audio signals from two or more different
3 audio sources; and
4 each set of one or more spatial parameters corresponds to a different audio source in the auditory
5 scene.

1 3. (canceled)

1 4. (original) The invention of claim 1, wherein the input audio signal is a mono signal.

1 5. (previously pending) The invention of claim 1, wherein the input audio signal
2 corresponds to a combination of two or more different source signals, wherein the two or more different
3 frequency bands are selected by comparing magnitudes of the two or more different source signals,
4 wherein, for each of the two or more different frequency bands, one of the source signals dominates the
5 other source signals.

1 6. (previously pending) The invention of claim 1, wherein the input audio signal
2 corresponds to a combination of left and right audio signals, wherein each different set of one or more
3 spatial parameters is generated by comparing the left and right audio signals in a corresponding
4 frequency band.

1 7. (original) The invention of claim 1, wherein step (a) comprises the step of dividing the
2 input audio signal into the plurality of different frequency bands based on information corresponding to
3 the different sets of one or more spatial parameters.

1 8. (previously pending) The invention of claim 1, wherein:
2 the input audio signal corresponds to a combination of audio signals from two or more different
3 audio sources; and
4 each set of one or more spatial parameters is applied to at least one frequency band in which the
5 input audio signal is dominated by a corresponding audio source in the auditory scene.

1 9. (original) The invention of claim 1, wherein each set of one or more spatial parameters
2 comprises one or more of an interaural level difference, an interaural time delay, and a head-related
3 transfer function.

1 10. (original) The invention of claim 1, wherein:
2 step (a) further comprises the step of converting the input audio signal from a time domain into a
3 frequency domain; and

4 step (b) further comprises the step of converting the two or more synthesized audio signals from
5 the frequency domain into the time domain.

1 11. (previously pending) The invention of claim 1, wherein the two or more synthesized
2 audio signals comprise left and right audio signals corresponding to the auditory scene.

1 12. (previously pending) The invention of claim 1, wherein the two or more synthesized
2 audio signals comprise three or more signals of a multi-channel audio signal corresponding to the
3 auditory scene.

1 13. (previously pending) The invention of claim 1, wherein:
2 the input audio signal is a mono signal;
3 each set of one or more spatial parameters corresponds to a different audio source in the auditory
4 scene;

5 step (a) comprises the steps of:

6 (1) converting the mono signal from a time domain into a frequency domain;
7 (2) dividing the converted mono signal into the plurality of different frequency
8 bands based on information corresponding to the sets of one or more spatial parameters;
9 each set of one or more spatial parameters is applied to at least one frequency band in which the
10 input audio signal is dominated by a corresponding audio source in the auditory scene;
11 each set of one or more spatial parameters comprises one or more of an interaural level
12 difference, an interaural time delay, and a head-related transfer function;
13 the two or more synthesized audio signals comprise left and right audio signals corresponding to
14 the auditory scene; and

15 step (b) further comprises the step of converting the left and right audio signals from the
16 frequency domain into the time domain.

1 14. (original) The invention of claim 13, wherein the mono signal corresponds to a
2 combination of two or more different mono source signals, wherein the two or more different frequency
3 bands are selected by comparing magnitudes of the two or more different mono source signals, wherein,
4 for each of the two or more different frequency bands, one of the mono source signals dominates the
5 other mono source signals.

1 15. (previously pending) The invention of claim 13, wherein the mono signal corresponds to
2 a combination of left and right audio signals, wherein each different set of one or more spatial
3 parameters is generated by comparing the left and right audio signals in a corresponding frequency band.

1 16. (currently amended) A machine-readable medium, having encoded thereon program
2 code, wherein, when the program code is executed by a machine, the machine implements a method for
3 synthesizing an auditory scene, comprising the steps of:

4 (a) dividing an input audio signal into a plurality of different frequency bands; and
5 (b) applying two or more different sets of one or more spatial parameters to two or more of
6 the different frequency bands in the input audio signal to generate two or more synthesized audio signals
7 of the auditory scene, wherein, for at least one of the sets of one or more spatial parameters, at least one
8 of the spatial parameters corresponds to a combination of two or more different audio sources in the
9 auditory scene that takes into account relative dominance of the two or more different audio sources in
10 the auditory scene.

1 17. (currently amended) An apparatus for synthesizing an auditory scene, comprising:
2 (a) means for dividing an input audio signal into a plurality of different frequency bands;
3 and
4 (b) means for applying two or more different sets of one or more spatial parameters to two
5 or more of the different frequency bands in the input audio signal to generate two or more synthesized
6 audio signals of the auditory scene, wherein, for at least one of the sets of one or more spatial parameters,
7 at least one of the spatial parameters corresponds to a combination of two or more different audio
8 sources in the auditory scene that takes into account relative dominance of the two or more different
9 audio sources in the auditory scene.

1 18. (currently amended) An apparatus for synthesizing an auditory scene, comprising:
2 (1) an auditory scene synthesizer configured to:
3 (a) divide an input audio signal into a plurality of different frequency bands; and
4 (b) apply two or more different sets of one or more spatial parameters to two or
5 more of the different frequency bands in the input audio signal to generate two or more synthesized
6 audio signals of the auditory scene; and
7 (2) one or more inverse time-frequency transformers configured to convert the two or more
8 synthesized audio signals from a frequency domain into a time domain, wherein, for at least one of the
9 sets of one or more spatial parameters, at least one of the spatial parameters corresponds to a
10 combination of two or more different audio sources in the auditory scene that takes into account relative
11 dominance of the two or more different audio sources in the auditory scene.

1 19. (currently amended) A method for processing two or more input audio signals,
2 comprising the steps of:
3 (a) converting the two or more input audio signals from a time domain into a frequency
4 domain;
5 (b) generating a set of one or more auditory scene parameters for each of two or more
6 different frequency bands in the two or more converted input audio signals; and
7 (c) combining the two or more input audio signals to generate a combined audio signal,
8 wherein, for at least one of the sets of one or more spatial parameters, at least one of the spatial
9 parameters corresponds to a combination of two or more different audio sources in an auditory scene that
10 takes into account relative dominance of the two or more different audio sources in the auditory scene.

1 20. (original) The invention of claim 19, wherein:
2 the two or more input audio signals are mono signals corresponding to different audio sources in
3 the auditory scene;
4 each set of one or more auditory scene parameters corresponds to an audio source that dominates
5 the other audio sources in the corresponding frequency band; and
6 the two or more input audio signals are combined in the time domain to generate the combined
7 audio signal.

1 21. (previously pending) The invention of claim 19, wherein:
2 the two or more input audio signals are left and right audio signals;
3 each set of one or more auditory scene parameters is generated by comparing the left and right
4 audio signals in the corresponding frequency band; and
5 further comprising the step of converting the combined audio signal from the frequency domain
6 into the time domain.

1 22. (currently amended) A machine-readable medium, having encoded thereon program
2 code, wherein, when the program code is executed by a machine, the machine implements a method for
3 processing two or more input audio signals, comprising the steps of:

- 4 (a) converting the two or more input audio signals from a time domain into a frequency
5 domain;
6 (b) generating a set of one or more auditory scene parameters for each of two or more
7 different frequency bands in the two or more converted input audio signals; and
8 (c) combining the two or more input audio signals to generate a combined audio signal,
9 wherein, for at least one of the sets of one or more spatial parameters, at least one of the spatial
10 parameters corresponds to a combination of two or more different audio sources in an auditory scene that
11 takes into account relative dominance of the two or more different audio sources in the auditory scene.

1 23. (currently amended) An apparatus for processing two or more input audio signals,
2 comprising:

- 3 (a) means for converting the two or more input audio signals from a time domain into a
4 frequency domain;
5 (b) means for generating a set of one or more auditory scene parameters for each of two or
6 more different frequency bands in the two or more converted input audio signals; and
7 (c) means for combining the two or more input audio signals to generate a combined audio
8 signal, wherein, for at least one of the sets of one or more spatial parameters, at least one of the spatial
9 parameters corresponds to a combination of two or more different audio sources in an auditory scene that
10 takes into account relative dominance of the two or more different audio sources in the auditory scene.

1 24. (currently amended) An apparatus for processing two or more input audio signals,
2 comprising:

- 3 (a) a time-frequency transformer configured to convert the two or more input audio signals
4 from a time domain into a frequency domain;
5 (b) an auditory scene parameter generator configured to generate a set of one or more
6 auditory scene parameters for each of two or more different frequency bands in the two or more
7 converted input audio signals; and
8 (c) a combiner configured to combine the two or more input audio signals to generate a
9 combined audio signal, wherein, for at least one of the sets of one or more spatial parameters, at least one
10 of the spatial parameters corresponds to a combination of two or more different audio sources in an
11 auditory scene that takes into account relative dominance of the two or more different audio sources in
12 the auditory scene.

1 25. (original) The invention of claim 24, wherein:

- 2 the two or more input audio signals are mono signals corresponding to different audio sources in
3 the auditory scene;
4 each set of one or more auditory scene parameters corresponds to an audio source that dominates
5 the other audio sources in the corresponding frequency band; and
6 the combiner operates in the time domain.

1 26. (previously pending) The invention of claim 24, wherein:

- 2 the two or more input audio signals are left and right audio signals;
3 each set of one or more auditory scene parameters is generated by comparing the left and right
4 audio signals in the corresponding frequency band; and
5 further comprising an inverse time-frequency transformer configured to convert the combined
6 audio signal from the frequency domain into the time domain.

1 27. (previously pending) The invention of claim 1, wherein, for each of the two or more
2 different frequency bands, the corresponding set of one or more spatial parameters is applied to the input
3 audio signal as if the input audio signal corresponded to a single audio source in the auditory scene.

1 28. (previously pending) The invention of claim 1, wherein the input audio signal
2 corresponds to a combination of three or more audio signals of a multi-channel signal, wherein each
3 different set of one or more spatial parameters is generated by comparing at least two of the audio signals
4 in a corresponding frequency band.

1 29. (previously pending) The invention of claim 1, further comprising decompressing a
2 compressed audio signal to generate the input audio signal.

1 30. (previously pending) The invention of claim 19, wherein each set of one or more
2 auditory scene parameters is generated as if the corresponding frequency band corresponded to a single
3 audio source in an auditory scene.

1 31. (previously pending) The invention of claim 19, wherein:
2 the two or more input audio signals are three or more audio signals of a multi-channel signal; and
3 each set of one or more auditory scene parameters is generated by comparing at least two of the
4 audio signals in the corresponding frequency band.

1 32. (previously pending) The invention of claim 19, further comprising compressing the
2 combined audio signal to generate a compressed audio signal.

1 33. (previously pending) The invention of claim 19, wherein the combined audio signal is
2 generated by performing auditory scene removal on the input audio signals in the frequency domain
3 based on the two or more sets of one or more auditory scene parameters.

1 34. (previously pending) The invention of claim 19, wherein the combined audio signal is
2 generated by averaging the input audio signals.

1 35. (currently amended) A bitstream comprising a combined audio signal and a plurality of
2 auditory scene parameters, wherein:
3 the combined audio signal is generated by combining two or more input audio signals; and
4 the auditory scene parameters are generated by:
5 converting the two or more input audio signals from a time domain into a frequency
6 domain; and
7 generating a set of one or more auditory scene parameters for each of two or more
8 different frequency bands in the two or more converted input audio signals, wherein, for at least one of
9 the sets of one or more spatial parameters, at least one of the spatial parameters corresponds to a
10 combination of two or more different audio sources in an auditory scene that takes into account relative
11 dominance of the two or more different audio sources in the auditory scene.

1 36. (new) A method for synthesizing an auditory scene, comprising the steps of:
2 (a) dividing an input audio signal into a plurality of different frequency bands; and
3 (b) applying two or more different sets of one or more spatial parameters to two or more of
4 the different frequency bands in the input audio signal to generate two or more synthesized audio signals
5 of the auditory scene, wherein the input audio signal corresponds to a combination of two or more
6 different source signals, wherein the two or more different frequency bands are selected by comparing

7 magnitudes of the two or more different source signals, wherein, for each of the two or more different
8 frequency bands, one of the source signals dominates the other source signals.

1 37. (new) A method for synthesizing an auditory scene, comprising the steps of:
2 (a) dividing an input audio signal into a plurality of different frequency bands; and
3 (b) applying two or more different sets of one or more spatial parameters to two or more of
4 the different frequency bands in the input audio signal to generate two or more synthesized audio signals
5 of the auditory scene, wherein the input audio signal corresponds to a combination of left and right audio
6 signals, wherein each different set of one or more spatial parameters is generated by comparing the left
7 and right audio signals in a corresponding frequency band.

1 38. (new) A method for synthesizing an auditory scene, comprising the steps of:
2 (a) dividing an input audio signal into a plurality of different frequency bands; and
3 (b) applying two or more different sets of one or more spatial parameters to two or more of
4 the different frequency bands in the input audio signal to generate two or more synthesized audio signals
5 of the auditory scene, wherein:
6 the input audio signal corresponds to a combination of audio signals from two or more different
7 audio sources; and
8 each set of one or more spatial parameters is applied to at least one frequency band in which the
9 input audio signal is dominated by a corresponding audio source in the auditory scene.

1 39. (new) The invention of claim 38, wherein:
2 the input audio signal is a mono signal;
3 each set of one or more spatial parameters corresponds to a different audio source in the auditory
4 scene;
5 step (a) comprises the steps of:
6 (1) converting the mono signal from a time domain into a frequency domain;
7 (2) dividing the converted mono signal into the plurality of different frequency
8 bands based on information corresponding to the sets of one or more spatial parameters;
9 each set of one or more spatial parameters comprises one or more of an interaural level
10 difference, an interaural time delay, and a head-related transfer function;
11 the two or more synthesized audio signals comprise left and right audio signals corresponding to
12 the auditory scene; and
13 step (b) further comprises the step of converting the left and right audio signals from the
14 frequency domain into the time domain.

1 40. (new) The invention of claim 39, wherein the mono signal corresponds to a combination
2 of two or more different mono source signals, wherein the two or more different frequency bands are
3 selected by comparing magnitudes of the two or more different mono source signals, wherein, for each of
4 the two or more different frequency bands, one of the mono source signals dominates the other mono
5 source signals.

1 41. (new) The invention of claim 39, wherein the mono signal corresponds to a combination
2 of left and right audio signals, wherein each different set of one or more spatial parameters is generated
3 by comparing the left and right audio signals in a corresponding frequency band.

1 42. (new) A method for synthesizing an auditory scene, comprising the steps of:
2 (a) dividing an input audio signal into a plurality of different frequency bands; and
3 (b) applying two or more different sets of one or more spatial parameters to two or more of
4 the different frequency bands in the input audio signal to generate two or more synthesized audio signals

5 of the auditory scene, wherein the input audio signal corresponds to a combination of three or more audio
6 signals of a multi-channel signal, wherein each different set of one or more spatial parameters is
7 generated by comparing at least two of the audio signals in a corresponding frequency band.

1 43. (new) A method for processing two or more input audio signals, comprising the steps
2 of:
3 (a) converting the two or more input audio signals from a time domain into a frequency
4 domain;
5 (b) generating a set of one or more auditory scene parameters for each of two or more
6 different frequency bands in the two or more converted input audio signals; and
7 (c) combining the two or more input audio signals to generate a combined audio signal,
8 wherein:

9 the two or more input audio signals are mono signals corresponding to different audio
10 sources in the auditory scene;

11 each set of one or more auditory scene parameters corresponds to an audio source that
12 dominates the other audio sources in the corresponding frequency band; and

13 the two or more input audio signals are combined in the time domain to generate the
14 combined audio signal.

1 44. (new) A method for processing two or more input audio signals, comprising the steps
2 of:
3 (a) converting the two or more input audio signals from a time domain into a frequency
4 domain;
5 (b) generating a set of one or more auditory scene parameters for each of two or more
6 different frequency bands in the two or more converted input audio signals; and
7 (c) combining the two or more input audio signals to generate a combined audio signal,
8 wherein:

9 the two or more input audio signals are left and right audio signals;

10 each set of one or more auditory scene parameters is generated by comparing the left and
11 right audio signals in the corresponding frequency band; and

12 further comprising the step of converting the combined audio signal from the frequency
13 domain into the time domain.

1 45. (new) A method for processing two or more input audio signals, comprising the steps
2 of:
3 (a) converting the two or more input audio signals from a time domain into a frequency
4 domain;
5 (b) generating a set of one or more auditory scene parameters for each of two or more
6 different frequency bands in the two or more converted input audio signals; and
7 (c) combining the two or more input audio signals to generate a combined audio signal,
8 wherein:
9 the two or more input audio signals are three or more audio signals of a multi-channel
10 signal; and
11 each set of one or more auditory scene parameters is generated by comparing at least two
12 of the audio signals in the corresponding frequency band.

1 46. (new) A method for processing two or more input audio signals, comprising the steps
2 of:
3 (a) converting the two or more input audio signals from a time domain into a frequency
4 domain;

5 (b) generating a set of one or more auditory scene parameters for each of two or more
6 different frequency bands in the two or more converted input audio signals; and
7 (c) combining the two or more input audio signals to generate a combined audio signal,
8 wherein the combined audio signal is generated by performing auditory scene removal on the input audio
9 signals in the frequency domain based on the two or more sets of one or more auditory scene parameters.

1 47. (new) A method for processing two or more input audio signals, comprising the steps
2 of:
3 (a) converting the two or more input audio signals from a time domain into a frequency
4 domain;
5 (b) generating a set of one or more auditory scene parameters for each of two or more
6 different frequency bands in the two or more converted input audio signals; and
7 (c) combining the two or more input audio signals to generate a combined audio signal,
8 wherein the combined audio signal is generated by averaging the input audio signals.

1 48. (new) An apparatus for processing two or more input audio signals, comprising:

2 (a) a time-frequency transformer configured to convert the two or more input audio signals

3 from a time domain into a frequency domain;

4 (b) an auditory scene parameter generator configured to generate a set of one or more

5 auditory scene parameters for each of two or more different frequency bands in the two or more

6 converted input audio signals; and

7 (c) a combiner configured to combine the two or more input audio signals to generate a

8 combined audio signal, wherein:

9 the two or more input audio signals are mono signals corresponding to different audio

0 sources in the auditory scene;

1 each set of one or more auditory scene parameters corresponds to an audio source that

2 dominates the other audio sources in the corresponding frequency band; and

3 the combiner operates in the time domain.

1 49. (new) An apparatus for processing two or more input audio signals, comprising:
2 (a) a time-frequency transformer configured to convert the two or more input audio signals
3 from a time domain into a frequency domain;
4 (b) an auditory scene parameter generator configured to generate a set of one or more
5 auditory scene parameters for each of two or more different frequency bands in the two or more
6 converted input audio signals; and
7 (c) a combiner configured to combine the two or more input audio signals to generate a
8 combined audio signal, wherein:
9 the two or more input audio signals are left and right audio signals;
10 each set of one or more auditory scene parameters is generated by comparing the left and
11 right audio signals in the corresponding frequency band; and
12 further comprising an inverse time-frequency transformer configured to convert the
13 combined audio signal from the frequency domain into the time domain.